Technical Implementation Committee Report 2010/11

Submitted by Wendy Thomas, Chair

Group members

- J. Gager, Metadata Technology, Technical Consultant
- Arofan Gregory, Metadata Technology , Technical Consultant
- Pascal Heus, Open Data Foundation (ODaF)
- Jeremy Iverson, Algenta Technologies, Invited Expert
- Dan Smith, Algenta Technologies, Invited Expert
- Wendy Thomas, University of Minnesota, Minnesota Population Center, Chair
- Joachim Wackerow, GESIS -- Leibniz Institute for the Social Sciences, Vice Chair

2.5 Update

After collecting specific requests for changes to support the IHSN Microdata Toolkit and reviewing the issues in transforming between DDI 2.1 and 3, we created a preliminary 2.5 version for internal review. This review process has taken longer than anticipated but DDI 2.5 will be released for review soon along with a new DDI editing tool created by Ben Youngdahl, Minnesota Population Center, which will be useful in evaluating 2.5 as well as future updates to DDI in general.

3.2 Update

TIC met in Minneapolis September 2010 to address the remaining 167 bugs as well as discuss process changes. We currently have approximate 35 remaining bugs which will be addressed in 3.2. The original goal was to provide a 3.2 version for review in Jan/Feb 2011. The bug list consisted of small individual bugs as well as sets of bugs related to the revision of the following areas:

- Identification and maintenance
- Complex question structures
- Response domains and representation
- Reusable content in variables
- Improving support for programming against control constructs
- Data relationships
- Improving reusability of Geographic Structure and Geographic Location
- Capturing variable and category statistics accurately
- Supporting controlled vocabularies
- Improved structure for capturing multi-lingual content and defining language

- Restructuring Organization Scheme
- Code scheme management
- Interaction of conceptual components, subject and keyword content

The most critical of these are the issues raised around identification and maintenance. There are two sets of requirements addressing technical implementation and business perspectives. TIC has spent considerable time defining these issues and trying to identify solutions that address both sets of requirements. This issue requires broader input from the DDI Expert and User community. Our decision has been to delay release of 3.2 until these issues are resolved.

Currently we are drafting a discussion piece to be distributed to a representative group of DDI users. Once we have feedback from this group we will have an in-depth discussion with a subset of this group and make a proposal on how identification within DDI should be structured. In the past we have examined specific aspects of this issue but are now coming to grips with the full set of issues. Currently the basic sub-issues include:

1) An object needs a single unchanging identifier
2) We need to be able to identify the context of an object as well as the object for purposes of documentation clarity for the user
3) We need to provide metadata creators and managers options in managing the creation and resolution process

Process Changes

In working through the above issues and dealing with the increase of bugs coming from DDI implementers, TIC has developed a number of new processes to improve communication with developers and increase the involvement of the Expert Committee and DDI users in the development and evaluation of proposed improvements and bug solutions.

Development site

In response to requests from the development community, TIC created a development subversion site. This site allows us to update schemas as bugs are resolved and make these development schemas available for evaluation by implementers. The new site supports early testing of bug resolutions. This means that schema versions released for formal review have had more extensive technical evaluation by interested implementers. Currently there are 13 individuals who have requested and received access to the development site (TIC member plus an additional 7 implementers). The site has been well received by implementers.

Proposed process for complex bug sets

A number of the issue areas noted above were brought to light by a number of related bugs. Resolving the bugs that were filed also brought to light related issues and resulted in the redesign or expansion of a number of features. In addition to bringing proposed solutions back to the individuals who filed the

original bugs, TIC would like to get feedback from Expert Committee members and DDI users who make use of these features. The approach we would like to implement is as follows.

TIC will write up the issues addressed by the proposed change as well as a brief discussion of how the proposed change resolves these issues. The proposed change will be provided and an example provided. TIC will then identify a sub-group of the Expert Committee and DDI user community who actively work with the feature and request feedback on the proposal. This will take place during the development process (prior to formal review). In order to do this we propose to initiate a survey of the Expert Committee and DDI users to identify individuals and their areas of expertise and DDI usage.

Plans for the results of SDI group

SDI completed its preliminary work early this year. Once we have completed version 3.2 we will take up the SDI proposals concerning sampling and survey development. The first step will be to prepare schema sections (new or additions to current schema) for testing. At that point we will need people working in this area that will be willing to test out new schema sections. The SDI is the first working group operating under a new process that included intensive involvement of a TIC member as they developed their recommendations. TIC is very pleased with the output of this group and the process that was followed in completing this work.